Title: An Introduction to Public Key Cryptography

Brief Overview:

Public key cryptography is a common means of securing information on the Internet. It is often used to protect credit card transmissions for on-line shopping. One of the earliest public key algorithms was the RSA algorithm, named for its inventors Ronald Rivest, Adi Shamir and Leonard Adelman. RSA is an important but often transparent part of Internet browsers such as Internet Explorer. In the lessons outlined below, students will learn how to encrypt and decrypt using RSA. In the process of learning RSA, students will become familiar with modular reduction of natural numbers and gain an understanding of some prime number theory.

NCTM 2000 Principles for School Mathematics:

- **Equity:** *Excellence in mathematics education requires equity - high expectations and strong support for all students.*
- **Curriculum:** *A curriculum is more than a collection of activities: it must be coherent, focused on important mathematics, and well articulated across the grades.*
- **Teaching:** *Effective mathematics teaching requires understanding what students know and need to learn and then challenging and supporting them to learn it well.*
- **Learning:** *Students must learn mathematics with understanding, actively building new knowledge from experience and prior knowledge.*
- **Assessment:** *Assessment should support the learning of important mathematics and furnish useful information to both teachers and students.*
- **Technology:** *Technology is essential in teaching and learning mathematics; it influences the mathematics that is taught and enhances students' learning.*

Links to NCTM 2000 Standards:

- **Content Standards**

Number and Operations
Students will use modular reduction of natural numbers to represent the mathematical processes they are examining.

Algebra
Students will use symbolic forms to develop mathematical constructs for modeling the processes of encryption and decryption.

- **Process Standards**

  Problem Solving
  Students will use logic skills and numerical data to encrypt and decrypt messages.

  Reasoning and Proof
  Students will use inductive reasoning to devise mathematically sound approaches for encrypting and decrypting messages.

  Communication
  Students will communicate their ideas using mathematical symbols to help construct their statements and arguments about how to model the encryption/decryption process.

  Connections
  Students will discover connections between algebra and numerical analysis.

  Representation
  Students will learn to represent the processes of encryption and decryption using mathematical equations.

Links to Maryland High School Mathematics Core Learning Units:

Functions and Algebra

- **1.1.1**
  Students will recognize, describe, and extend patterns and functional relationships that are expressed numerically and algebraically.

- **1.1.3**
  Students will add, subtract, multiply, and divide algebraic expressions.

Geometry, Measurement, and Reasoning

- **2.2.3**
  Students will identify and use inductive and deductive reasoning.

Grade/Level:

Grades 11-12

Duration/Length:

Three 45-minute periods

Prerequisite Knowledge:

Students should have working knowledge of the following skills:

- Algebra
- Basic arithmetic

Student Outcomes:

Students will:

- gain a basic understanding of prime number theory.
- understand modular reduction with natural numbers.
- understand how to find the greatest common denominator of natural numbers in mod *n*.
- learn the basic principles of encryption and decryption.

Materials/Resources/Printed Materials:

- Lewand, R.E., (2000). Cryptological Mathematics. Washington, DC: The Mathematical Association of America.
- Sinkov, A., (1966). Elementary Cryptanalysis: A Mathematical Approach. Washington, DC: The Mathematical Association of America.
- Pencil
- Paper
- Advanced scientific calculator

Development/Procedures:

Lesson 1: An Introduction to Prime Number Theory

The RSA encryption algorithm builds encryption and decryption keys from two extremely large prime numbers (in a manner that will be discussed in Lesson 3). As a result, prime number theory is the foundation for RSA. This lesson will start from the definition of prime numbers and proceed to introduce the basic number theory concepts that are needed for RSA.

Definitions: A *prime number* is a positive integer greater than one that is divisible by no positive integers other than one and itself. A positive integer which is not prime and which is not equal to one is called a *composite number*.

The importance of prime numbers can be seen in the following result.

The Fundamental Theorem of Arithmetic: Every positive integer greater than one can be written uniquely as a product of primes, with the prime factors in the product written in order of nondecreasing size.

This result tells us that prime numbers are the basic building blocks of the integers.

Furthermore, it allows us to create an encryption algorithm using prime numbers. For RSA encryption, we start with two primes *p* and *q* and use the number $n = pq$. In order for decryption to be possible, there must be a way to get back to *p* and *q* from *n*. The Fundamental Theorem guarantees that *n* has a unique decomposition into primes and thus a factorization of *n* must yield only *p* and q. The strength of the RSA algorithm depends on the size of *p* and *q* and the resulting difficulty an adversary would have in factoring the product not knowing *p* and *q*.

Although the Fundamental Theorem indicates that we can build an encryption algorithm using primes, it is not clear that we can find large enough primes to develop a secure algorithm. Therefore, we next show that given any collection of primes we can always find another prime number not in that collection.

Let $S = \{P_1, P_2, \ldots, P_N\}$ be a finite collection of prime numbers and consider the number $P = (P_1 \times P_2 \times \ldots \times P_N) + 1$. By the Fundamental Theorem of Algebra, either P is prime or it is a product of prime numbers. If P is prime, then it is not an element in S since it is bigger than any element of S. If P is a product of prime numbers, then there is a prime divisor of P not in S since no element of S divides P. Thus, there is an infinite number of prime numbers, and given the first N prime numbers we can, in theory, find a prime larger than any of these.

However, in practice, in is not easy to find prime numbers. For the final part of this lesson, we discuss the sieve of Eratosthenes, one of the earliest methods of finding prime numbers and which serves as a foundation for the modern methods of finding prime numbers.

To illustrate this method for finding prime numbers, we illustrate how to find all prime numbers less than 100. The positive integers between 1 and 100 are written in a ten by ten grid. We first cross out all numbers other than two in the grid which are multiples of two. Then, all remaining integers other than three, which are multiples of three, are crossed out. Similarly, all remaining multiples of 5 and 7 are in turn crossed out. The numbers in the grid which have not been crossed out are one and all the prime numbers less than one hundred. Note that a composite integer has no prime factors larger than its square root. So, in our example, it was not necessary to worry about crossing out multiples of numbers larger than ten.

The students now have enough information to answer problems 1 and 2 of the handout.

<u>Lesson 2:</u> Modular Arithmetic and Greatest Common Denominator

Modular arithmetic is a useful tool in computer science, mathematics and cryptography. In general terms, the modulus (mod) is the remainder of a division problem. For example:

$$73 \bmod 10 = 3 \text{ because } 73 = 7 * 10 + 3$$

Written mathematically,

$$a \bmod n = b \text{ where } a = k * n + b$$

Since k can be any integer, this equation has an infinite number of solutions. In order to limit our answer, b is chosen such that $0 \leq b < n$.

This is a useful concept in cryptology. The English alphabet only has 26 letters but numbers are infinite. If we represent the letters by the numbers 0 – 25, where a=0, b=1, c=2, …, y=24, z=25, then we can reduce any number to a value within the range of letters by setting the modulus to 26.

Another useful concept is modular congruency. Two numbers a and b are congruent modulo a number n if their difference (a - b) is a multiple of n. Formally, we express congruence in the following way.

$$a \equiv b \ (mod\ n)$$

Here are some examples to illustrate this concept.

$13 \equiv 5 \bmod 8$ since $13 - 5 = 8$ is a multiple of 8.
$5 \equiv 26 \bmod 7$ since $5 - 26 = -21$ is a multiple of 7.

The final concept that we will need in order to understand RSA encryption is how to find the greatest common denominator (gcd) of two numbers. The gcd of two numbers, a and b, is the largest number that divides both a and b. For example:

gcd (18, 48) = 6
gcd (17, 28) = 1

The easiest way to determine the gcd of two numbers is to perform the prime factorization of both numbers. Then multiply the common factors together to obtain the greatest number that can divide both of the original numbers. In our previous examples

18 = 1 * 2 * 3 * 3
48 = 1 * 2 * 2 * 2 * 2 * 3
17 = 1 * 17
28 = 1 * 2 * 2 * 7

In both of these factorizations, we see that 18 and 48 have one ‘1’, one ‘2’, and one ‘3’ in common. Therefore, the gcd (18, 48) = 1 * 2 * 3 = 6. In the case of 17 and 28, the only common factor they share is 1. Therefore, the gcd (17, 28) = 1. In this case, we say that these two numbers are relatively prime to one another. Many cryptologic techniques, including RSA, require the keys to be relatively prime to one another.

The students now have enough information to answer problems 3-5 of the handout.

Lesson 3: Encrypting and Decrypting Messages Using RSA

RSA is considered an asymmetric cryptographic system. In such systems, the sender and receiver do not share a common key; rather each participant has a *public key* **E** (for enciphering) and a *private key* **D** (for deciphering). The public and private keys of the sender do not relate to the public and private keys of the receiver of the message. For this reason, the public keys can be published for anyone to see. Here is an example to illustrate this concept.

Suppose Alice wants to send Bob an encrypted message *m*. Alice looks up Bob's public key E_B in a directory and enciphers *m* using this key – the result is a ciphertext message $E_B(m)$. Upon receiving the ciphertext, Bob deciphers it using his private key D_B and arrives at a plaintext message $D_B(E_B(m))$. If the system is correctly designed, $D_B(E_B(m)) = m$ and so Bob now can read Alice's original plaintext message.

While this process sounds simple, it is quite secure if the keys are chosen properly. The keys are derived as follows:

1. For each participant, a Key Center selects two (preferably extremely large) prime numbers: p and q.
2. The Center forms the product $n = pq$.
3. The Center computes $\phi(n) = (p-1)(q-1)$.
4. The Center selects any integer e with the property that $\gcd(e, \phi(n)) = 1$.
5. The Center finds an integer d with $ed \equiv 1 \bmod \phi(n)$.
6. The Center issues to the participant the private key (d) and the public keys (e and n). Note: The participant has no idea what the values of p and q are!

The security of the encryption depends on the difficulty of finding p and q. Here is a simplified example to illustrate encryption using the RSA algorithm:

Given the modulus $n = 55$ and encryption exponent $e = 3$, encrypt the numbers 7 and 27.

To encrypt, we raise the numerical representation of the plaintext to the power given by the encryption exponent and then reduce mod *n*. Since $7^3 = 343 = 6 * 55 + 13$, we have $7^3 \bmod 55 = 13$. So the plain text 7 is converted to the cipher text 13. Since $27^3 = 19683 = 357 * 55 + 48$, we have $27^3 \bmod 55 = 48$. So the plain text 27 is converted to the cipher text 48.

Decryption works just as simply as encryption if you have the decryption exponent *d*, which is chosen so that $ed \equiv 1 \bmod (p-1)(q-1)$. If the person decrypting the message is the intended recipient, he or she will possess d as their private key. If the message was intercepted, the person decrypting the message must first find *d*. We will proceed as if we had intercepted the message.

Consider the modulus 55 and the encryption exponent 3 (this is the intended recipient's public, published key). In the case of modulus 55, $p = 5$ and $q = 11$ since $5 * 11 = 55$ and both 5 and 11 are prime. We must have $3d \equiv 1 \bmod (5-1)(11-1)$ or $3d \equiv 1 \bmod 40$. Since $3 * 27 = 81 = 2 * 40 + 1$, we have $3 * 27 \equiv 1 \bmod 40$ and thus $d = 27$. (Note how easy this is when p and q are not extremely large.) Now that we have the decryption key, we can decrypt the numbers 13 and 48 (our cipher from above).

Since $13^{27} \bmod 55 = 7$, the cipher text 13 corresponds to the plain text 7.
Since $48^{27} \bmod 55 = 27$, the cipher text 48 corresponds to the plain text 27.

Let's work through an example with an actual alphabetic message.

Encrypt ALABAMA using the public key with modulus 319 and encryption exponent 3.

First convert the letters to numerical equivalents. The numerical equivalents could be the ascii code of the letter, the position of the letter in the alphabet, or any other similar scheme. We will use the position of the letter in the alphabet, starting with a = 1.

```
A  L  A  B  A  M  A
01 12 01 02 01 13 01
```

Group these numbers into blocks, adding a 23 (X) to fill the last block:

Plain Text: 0112 0102 0113 0123

Plain	Cipher
0112	$112^3 \bmod 319 = 52$
0102	$102^3 \bmod 319 = 214$
0113	$113^3 \bmod 319 = 60$
0123	$123^3 \bmod 319 = 140$

Therefore, the cipher text is 0052 0214 0060 0140.

The students now have enough information to answer problem 6 of the handout.

Assessment:

Student progress can be measured by defining an appropriate scoring rubric based on the student's performance on the attached handout. One suggested rubric is listed below and is based on the point values assigned to each problem in the handout.

18 – 20 Work is complete and correct
16 – 17 Work is almost complete and correct. Some minor errors may be observed.
14 – 15 Work is fairly complete and correct and indicates a general understanding of concepts. Noticeable errors are evident.
12 – 13 Some work is complete and correct. Significant errors that indicate a minimal understanding of the concepts are evident.

0 – 11 Work is wrong. There is no evidence that the student has grasped the concepts.

The evaluation of the handout should be based on:
(1) Description of mathematical concepts.
(2) Use and manipulation of mathematical equations
(3) Use of logical reasoning.

Extension/Follow Up:

Students could study other systems of encryption and decryption on their own. Many classical cryptologic systems can be found in Elementary Cryptanalysis. This book also explores how to break these systems.

Authors:

Mathematics Education Partnership Program
National Security Agency
Fort Meade, MD

HANDOUT

1. Let $S = \{ p_1, p_2, \ldots, p_N\}$ be the set containing the first N prime numbers and let $P = (p_1 * p_2 * \ldots * p_N) + 1$. Find a prime divisor of P that is not in S when

   a) $N = 4$

   b) $N = 6$

2. Find all prime numbers less than 250.

3. Reduce the following numbers to values within the range (0, n), where n is the modulus.

   a) 52 (mod 25)
   b) 79 (mod 7)
   c) 19 (mod 3)
   d) 589 (mod 17)
   e) 3998 (mod 56)
   f) 70 (mod 26)

4. State whether the following are true or false.

   a) $14 \equiv 5 \pmod 9$
   b) $4 \equiv 16 \pmod{12}$
   c) $7 \equiv 3 \pmod{10}$
   d) $-3 \equiv 5 \pmod{18}$
   e) $3 \equiv 9 \pmod{12}$
   f) $90 \equiv 9 \pmod{10}$

5. Find the gcd.

   a) gcd (15, 5)
   b) gcd (89, 46)
   c) gcd (87, 190)
   d) gcd (54, 7)
   e) gcd (41, 101)
   f) gcd (890, 21)

6. Given the public key with modulus 2911 and encryption exponent 1867, decrypt the following RSA-encrypted message:

   0518 0753 0875 1618 1126 0615 1613 1966 2062 2412 1173 0030 2756 0308 2524 0376

   0918 2528 2169 1359 0534 2116 0126 1606 0330

HANDOUT ANSWERS

1. a) S = { 2, 3, 5, 7}
   P = 211
   prime divisor = 211

   b) S = {2, 3, 5, 7, 11, 13}
   P = 30031
   prime divisor = 59 and 509

2. {2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97, 101, 103, 107, 109, 113, 127, 131, 137, 139, 149, 151, 157, 163, 167, 173, 179, 181, 191, 193, 197, 199, 211, 223, 227, 229, 233, 239, 241}

3. a) 52 mod 25 = 2
   b) 79 mod 7 = 2
   c) 19 mod 3 = 1
   d) 589 mod 17 = 11
   e) 3998 mod 56 = 22
   f) 70 mod 26 = 18

4. a) $14 \equiv 5 \bmod 9$ TRUE
   b) $4 \equiv 16 \bmod 12$ TRUE
   c) $7 \equiv 3 \bmod 10$ FALSE
   d) $-3 \equiv 5 \bmod 18$ FALSE
   e) $3 \equiv 9 \bmod 12$ FALSE
   f) $90 \equiv 9 \bmod 10$ TRUE

5. a) gcd (15, 5) = 5
   b) gcd (89, 46) = 1
   c) gcd (87, 190) = 1
   d) gcd (54, 7) = 1
   e) gcd (41, 101) = 1
   f) gcd (890, 21) = 1

6. CONGRATULATIONS STOP YOU HAVE DECRYPTED MY SECRET MESSAGE